

#8

12/01/92  
12:38:01  
S4120.raw

Patent Application US/07/864,692

## SEQUENCE LISTING



## (1) GENERAL INFORMATION:

(i) APPLICANT: Israel, David  
Wolfman, Neil M.

(ii) TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein  
Heterodimers, Compositions and Methods of Use.

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Legal Affairs, Genetics Institute, Inc.  
(B) STREET: 87 CambridgePark Drive  
(C) CITY: Cambridge  
(D) STATE: MA  
(E) COUNTRY: USA  
(F) ZIP: 02140-2387

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Tape  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US  
(B) FILING DATE:  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Kapinos, Ellen J.  
(B) REGISTRATION NUMBER: 32,245  
(C) REFERENCE/DOCKET NUMBER: GI-5192B

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617-876-1170  
(B) TELEFAX: 617-876-5851

## (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1607 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

## Patent Application US/07/864,692

53 ?  
54 (ix) FEATURE:  
55 (A) NAME/KEY: CDS  
56 (B) LOCATION: 356..1543  
57  
58  
59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:  
60  
61 GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGACTT CTTGAACTTG CAGGGAGAAT 60  
62  
63 AACTTGCGCA CCCCACTTTG CGCCGGTGCC TTTGCCCCAG CGGAGCCTGC TTCGCCATCT 120  
64  
65 CCGAGCCCCA CCGCCCCTCC ACTCCTCGGC CTTGCCCCGAC ACTGAGACGC TGTTCACAGC 180  
66  
67 GTGAAAAGAG AGACTGCGCG GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG 240  
68  
69 GACATTCCGGT CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT TCCATGTGGA CGCTCTTTCA 300  
70  
71 ATGGACGTGT CCCC GCGTGC TTCTTAGACG GACTGCGGTC TCCTAAAGGT CGACC ATG 358  
72 Met  
73 1  
74  
75 GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC CTC 406  
76 Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val Leu  
77 5 10 15  
78  
79 CTG GGC GGC GCG GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG TTC 454  
80 Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys Phe  
81 20 25 30  
82  
83 GCG GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG GTC 502  
84 Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val  
85 35 40 45  
86  
87 CTG AGC GAG TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG 550  
88 Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys Gln  
89 50 55 60 65  
90  
91 AGA CCC ACC CCC AGC AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA GAC 598  
92 Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu Asp  
93 70 75 80  
94  
95 CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC CAC 646  
96 Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp His  
97 85 90 95  
98  
99 CGG TTG GAG AGG GCA GCC AGC CGA GCC AAC ACT GTG CGC AGC TTC CAC 694  
100 Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe His  
101 100 105 110  
102  
103 CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG AGT GGG AAA ACA ACC 742  
104 His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr Thr

## Patent Application US/07/864,692

	115	120	125	
105				
106				
107	CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG GAG TTT ATC			790
108	Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe Ile			
109	130	135	140	145
110				
111	ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT TTA			838
112	Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala Leu			
113		150	155	160
114				
115	GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC ATA			886
116	Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile Ile			
117		165	170	175
118				
119	AAA CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGA CTT TTG GAC			934
120	Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu Asp			
121		180	185	190
122				
123	ACC AGG TTG GTG AAT CAG AAT GCA AGC AGG TGG GAA ACT TTT GAT GTC			982
124	Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Thr Phe Asp Val			
125		195	200	205
126				
127	ACC CCC GCT GTG ATG CGG TGG ACT GCA CAG GGA CAC GCC AAC CAT GGA			1030
128	Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His Gly			
129	210	215	220	225
130				
131	TTC GTG GTG GAA GTG GCC CAC TTG GAG GAG AAA CAA GGT GTC TCC AAG			1078
132	Phe Val Val Glu Val Ala His Leu Glu Lys Gln Gly Val Ser Lys			
133		230	235	240
134				
135	AGA CAT GTT AGG ATA AGC AGG TCT TTG CAC CAA GAT GAA CAC AGC TGG			1126
136	Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser Trp			
137		245	250	255
138				
139	TCA CAG ATA AGG CCA TTG CTA GTA ACT TTT GGC CAT GAT GGA AAA GGG			1174
140	Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys Gly			
141		260	265	270
142				
143	CAT CCT CTC CAC AAA AGA GAA AAA CGT CAA GCC AAA CAC AAA CAG CGG			1222
144	His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys His Lys Gln Arg			
145		275	280	285
146				
147	AAA CGC CTT AAG TCC AGC TGT AAG AGA CAC CCT TTG TAC GTG GAC TTC			1270
148	Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp Phe			
149	290	295	300	305
150				
151	AGT GAC GTG GGG TGG AAT GAC TGG ATT GTG GCT CCC CCG GGG TAT CAC			1318
152	Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr His			
153		310	315	320
154				
155	GCC TTT TAC TGC CAC GGA GAA TGC CCT TTT CCT CTG GCT GAT CAT CTG			1366
156	Ala Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His Leu			

## Patent Application US/07/864,692

157 i 325 330 335  
158  
159 AAC TCC ACT AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT GTT AAC 1414  
160 Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn  
161 340 345 350  
162  
163 TCT AAG ATT CCT AAG GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC 1462  
164 Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile  
165 355 360 365  
166  
167 TCG ATG CTG TAC CTT GAC GAG AAT GAA AAG GTT GTA TTA AAG AAC TAT 1510  
168 Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn Tyr  
169 370 375 380 385  
170  
171 CAG GAC ATG GTT GTG GAG GGT TGT GGG TGT CGC TAGTACAGCA AAATTAAATA 1563  
172 Gln Asp Met Val Val Glu Gly Cys Gly Cys Arg  
173 390 395  
174  
175 CATAAATATA TATATATATA TATATTTTAG AAAAAAGAAA AAAA 1607  
176  
177  
178 (2) INFORMATION FOR SEQ ID NO:2:  
179  
180 (i) SEQUENCE CHARACTERISTICS:  
181 (A) LENGTH: 396 amino acids  
182 (B) TYPE: amino acid  
183 (D) TOPOLOGY: linear  
184  
185 (ii) MOLECULE TYPE: protein  
186  
187 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:  
188  
189 Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val  
190 1 5 10 15  
191  
192 Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys  
193 20 25 30  
194  
195 Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu  
196 35 40 45  
197  
198 Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys  
199 50 55 60  
200  
201 Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu  
202 65 70 75 80  
203  
204 Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp  
205 85 90 95  
206  
207 His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe  
208 100 105 110

## Patent Application US/07/864,692

209  
210 His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr  
211 115 120 125  
212  
213 Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe  
214 130 135 140  
215  
216 Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala  
217 145 150 155 160  
218  
219 Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile  
220 165 170 175  
221  
222 Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu  
223 180 185 190  
224  
225 Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Thr Phe Asp  
226 195 200 205  
227  
228 Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His  
229 210 215 220  
230  
231 Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser  
232 225 230 235 240  
233  
234 Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser  
235 245 250 255  
236  
237 Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys  
238 260 265 270  
239  
240 Gly His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys His Lys Gln  
241 275 280 285  
242  
243 Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp  
244 290 295 300  
245  
246 Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr  
247 305 310 315 320  
248  
249 His Ala Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His  
250 325 330 335  
251  
252 Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val  
253 340 345 350  
254  
255 Asn Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala  
256 355 360 365  
257  
258 Ile Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn  
259 370 375 380  
260

## Patent Application US/07/864,692

261 Tyr Gln Asp Met Val Val Glu Gly Cys Gly Cys Arg  
262 385 390 395

263

264 (2) INFORMATION FOR SEQ ID NO:3:

265

266 (i) SEQUENCE CHARACTERISTICS:

267 (A) LENGTH: 1954 base pairs

268 (B) TYPE: nucleic acid

269 (C) STRANDEDNESS: double

270 (D) TOPOLOGY: unknown

271

272 (ii) MOLECULE TYPE: DNA (genomic)

273

274

275 (ix) FEATURE:

276 (A) NAME/KEY: CDS

277 (B) LOCATION: 403..1626

278

279

280 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

281

282 CTCTAGAGGG CAGAGGAGGA GGGAGGGAGG GAAGGAGCGC GGAGCCCGGC CCGGAAGCTA 60

283

284 GGTGAGTGTG GCATCCGAGC TGAGGGACGC GAGCCTGAGA CGCCGCTGCT GCTCCGGCTG 120

285

286 AGTATCTAGC TTGTCTCCCC GATGGGATTC CCGTCCAAGC TATCTCGAGC CTGCAGCGCC 180

287

288 ACAGTCCCCG GCCCTCGCCC AGGTTCACTG CAACCGTTCA GAGGTCCCCA GGAGCTGCTG 240

289

290 CTGGCGAGCC CGCTACTGCA GGGACCTATG GAGCCATTCC GTAGTGCCAT CCCGAGCAAC 300

291

292 GCACTGCTGC AGCTTCCCTG AGCCTTTCCA GCAAGTTTGT TCAAGATTGG CTGTCAAGAA 360

293

294 TCATGGACTG TTATTATATG CCTTGTTTTT TGTCAAGACA CC ATG ATT CCT GGT 414

295

296

297

298 AAC CGA ATG CTG ATG GTC GTT TTA TTA TGC CAA GTC CTG CTA GGA GGC 462

299

299 Asn Arg Met Leu Met Val Val Leu Leu Cys Gln Val Leu Leu Gly Gly

300

300 5 10 15 20 510

301

302 GCG AGC CAT GCT AGT TTG ATA CCT GAG ACG GGG AAG AAA AAA GTC GCC

303

303 Ala Ser His Ala Ser Leu Ile Pro Glu Thr Gly Lys Lys Lys Val Ala

304

304 25 30 35

305

306 GAG ATT CAG GGC CAC GCG GGA GGA CGC CGC TCA GGG CAG AGC CAT GAG 558

307

307 Glu Ile Gln Gly His Ala Gly Gly Arg Arg Ser Gly Gln Ser His Glu

308

308 40 45 50

309

310 CTC CTG CGG GAC TTC GAG GCG ACA CTT CTG CAG ATG TTT GGG CTG CGC 606

311

311 Leu Leu Arg Asp Phe Glu Ala Thr Leu Leu Gln Met Phe Gly Leu Arg

312

312 55 60 65

## Patent Application US/07/864,692

```

313
314 CGC CGC CCG CAG CCT AGC AAG AGT GCC GTC ATT CCG GAC TAC ATG CGG      654
315 Arg Arg Pro Gln Pro Ser Lys Ser Ala Val Ile Pro Asp Tyr Met Arg
316      70                      75                      80
317
318 GAT CTT TAC CGG CTT CAG TCT GGG GAG GAG GAG GAA GAG CAG ATC CAC      702
319 Asp Leu Tyr Arg Leu Gln Ser Gly Glu Glu Glu Glu Glu Gln Ile His
320      85                      90                      95                      100
321
322 AGC ACT GGT CTT GAG TAT CCT GAG CGC CCG GCC AGC CGG GCC AAC ACC      750
323 Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala Ser Arg Ala Asn Thr
324                      105                      110                      115
325
326 GTG AGG AGC TTC CAC CAC GAA GAA CAT CTG GAG AAC ATC CCA GGG ACC      798
327 Val Arg Ser Phe His His Glu Glu His Leu Glu Asn Ile Pro Gly Thr
328                      120                      125                      130
329
330 AGT GAA AAC TCT GCT TTT CGT TTC CTC TTT AAC CTC AGC AGC ATC CCT      846
331 Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu Ser Ser Ile Pro
332      135                      140                      145
333
334 GAG AAC GAG GTG ATC TCC TCT GCA GAG CTT CGG CTC TTC CGG GAG CAG      894
335 Glu Asn Glu Val Ile Ser Ser Ala Glu Leu Arg Leu Phe Arg Glu Gln
336      150                      155                      160
337
338 GTG GAC CAG GGC CCT GAT TGG GAA AGG GGC TTC CAC CGT ATA AAC ATT      942
339 Val Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His Arg Ile Asn Ile
340      165                      170                      175                      180
341
342 TAT GAG GTT ATG AAG CCC CCA GCA GAA GTG GTG CCT GGG CAC CTC ATC      990
343 Tyr Glu Val Met Lys Pro Pro Ala Glu Val Val Pro Gly His Leu Ile
344                      185                      190                      195
345
346 ACA CGA CTA CTG GAC ACG AGA CTG GTC CAC CAC AAT GTG ACA CGG TGG      1038
347 Thr Arg Leu Leu Asp Thr Arg Leu Val His His Asn Val Thr Arg Trp
348      200                      205                      210
349
350 GAA ACT TTT GAT GTG AGC CCT GCG GTC CTT CGC TGG ACC CGG GAG AAG      1086
351 Glu Thr Phe Asp Val Ser Pro Ala Val Leu Arg Trp Thr Arg Glu Lys
352      215                      220                      225
353
354 CAG CCA AAC TAT GGG CTA GCC ATT GAG GTG ACT CAC CTC CAT CAG ACT      1134
355 Gln Pro Asn Tyr Gly Leu Ala Ile Glu Val Thr His Leu His Gln Thr
356      230                      235                      240
357
358 CGG ACC CAC CAG GGC CAG CAT GTC AGG ATT AGC CGA TCG TTA CCT CAA      1182
359 Arg Thr His Gln Gly Gln His Val Arg Ile Ser Arg Ser Leu Pro Gln
360      245                      250                      255                      260
361
362 GGG AGT GGG AAT TGG GCC CAG CTC CGG CCC CTC CTG GTC ACC TTT GGC      1230
363 Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu Val Thr Phe Gly
364      265                      270                      275

```

## Patent Application US/07/864,692

365  
366 CAT GAT GGC CGG GGC CAT GCC TTG ACC CGA CGC CGG AGG GCC AAG CGT 1278  
367 His Asp Gly Arg Gly His Ala Leu Thr Arg Arg Arg Arg Ala Lys Arg  
368 280 285 290  
369  
370 AGC CCT AAG CAT CAC TCA CAG CGG GCC AGG AAG AAG AAT AAG AAC TGC 1326  
371 Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys  
372 295 300 305  
373  
374 CGG CGC CAC TCG CTC TAT GTG GAC TTC AGC GAT GTG GGC TGG AAT GAC 1374  
375 Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp  
376 310 315 320  
377  
378 TGG ATT GTG GCC CCA CCA GGC TAC CAG GCC TTC TAC TGC CAT GGG GAC 1422  
379 Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly Asp  
380 325 330 335 340  
381  
382 TGC CCC TTT CCA CTG GCT GAC CAC CTC AAC TCA ACC AAC CAT GCC ATT 1470  
383 Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala Ile  
384 345 350 355  
385  
386 GTG CAG ACC CTG GTC AAT TCT GTC AAT TCC AGT ATC CCC AAA GCC TGT 1518  
387 Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile Pro Lys Ala Cys  
388 360 365 370  
389  
390 TGT GTG CCC ACT GAA CTG AGT GCC ATC TCC ATG CTG TAC CTG GAT GAG 1566  
391 Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu  
392 375 380 385  
393  
394 TAT GAT AAG GTG GTA CTG AAA AAT TAT CAG GAG ATG GTA GTA GAG GGA 1614  
395 Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu Met Val Val Glu Gly  
396 390 395 400  
397  
398 TGT GGG TGC CGC TGAGATCAGG CAGTCCTTGA GGATAGACAG ATATACACAC 1666  
399 Cys Gly Cys Arg  
400 405  
401  
402 CACACACACA CACCACATAC ACCACACACA CACGTTCCCA TCCACTCACC CACACACTAC 1726  
403  
404 ACAGACTGCT TCCTTATAGC TGGACTTTTA TTTAAAAAAA AAAAAAAAAA AATGGAAAAA 1786  
405  
406 ATCCCTAAAC ATTACCTTG ACCTTATTTA TGACTTTACG TGCAAATGTT TTGACCATAT 1846  
407  
408 TGATCATATA TTTTGACAAA ATATATTTAT AACTACGTAT TAAAAGAAAA AAATAAAATG 1906  
409  
410 AGTCATTATT TTAAAAAAA AAAAAAACT CTAGAGTCGA CGGAATTC 1954  
411  
412  
413 (2) INFORMATION FOR SEQ ID NO:4:  
414  
415 (i) SEQUENCE CHARACTERISTICS:  
416 (A) LENGTH: 408 amino acids



## Patent Application US/07/864,692

417 (B) TYPE: amino acid  
418 (D) TOPOLOGY: linear  
419  
420 (ii) MOLECULE TYPE: protein  
421  
422 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:  
423  
424 Met Ile Pro Gly Asn Arg Met Leu Met Val Val Leu Leu Cys Gln Val  
425 1 5 10 15  
426  
427 Leu Leu Gly Gly Ala Ser His Ala Ser Leu Ile Pro Glu Thr Gly Lys  
428 20 25 30  
429  
430 Lys Lys Val Ala Glu Ile Gln Gly His Ala Gly Gly Arg Arg Ser Gly  
431 35 40 45  
432  
433 Gln Ser His Glu Leu Leu Arg Asp Phe Glu Ala Thr Leu Leu Gln Met  
434 50 55 60  
435  
436 Phe Gly Leu Arg Arg Arg Pro Gln Pro Ser Lys Ser Ala Val Ile Pro  
437 65 70 75 80  
438  
439 Asp Tyr Met Arg Asp Leu Tyr Arg Leu Gln Ser Gly Glu Glu Glu Glu  
440 85 90 95  
441  
442 Glu Gln Ile His Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala Ser  
443 100 105 110  
444  
445 Arg Ala Asn Thr Val Arg Ser Phe His His Glu Glu His Leu Glu Asn  
446 115 120 125  
447  
448 Ile Pro Gly Thr Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu  
449 130 135 140  
450  
451 Ser Ser Ile Pro Glu Asn Glu Val Ile Ser Ser Ala Glu Leu Arg Leu  
452 145 150 155 160  
453  
454 Phe Arg Glu Gln Val Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His  
455 165 170 175  
456  
457 Arg Ile Asn Ile Tyr Glu Val Met Lys Pro Pro Ala Glu Val Val Pro  
458 180 185 190  
459  
460 Gly His Leu Ile Thr Arg Leu Leu Asp Thr Arg Leu Val His His Asn  
461 195 200 205  
462  
463 Val Thr Arg Trp Glu Thr Phe Asp Val Ser Pro Ala Val Leu Arg Trp  
464 210 215 220  
465  
466 Thr Arg Glu Lys Gln Pro Asn Tyr Gly Leu Ala Ile Glu Val Thr His  
467 225 230 235 240  
468

## Patent Application US/07/864,692

469 Leu His Gln Thr Arg Thr His Gln Gly Gln His Val Arg Ile Ser Arg  
470 245 250 255  
471  
472 Ser Leu Pro Gln Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu  
473 260 265 270  
474  
475 Val Thr Phe Gly His Asp Gly Arg Gly His Ala Leu Thr Arg Arg Arg  
476 275 280 285  
477  
478 Arg Ala Lys Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys  
479 290 295 300  
480  
481 Asn Lys Asn Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val  
482 305 310 315 320  
483  
484 Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr  
485 325 330 335  
486  
487 Cys His Gly Asp Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr  
488 340 345 350  
489  
490 Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile  
491 355 360 365  
492  
493 Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu  
494 370 375 380  
495  
496 Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu Met  
497 385 390 395 400  
498  
499 Val Val Glu Gly Cys Gly Cys Arg  
500 405  
501

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1448 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 97..1389

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

520 GTGACCGAGC GGCGCGGACG GCCGCCTGCC CCCTCTGCCA CCTGGGGCGG TCGGGGCCCCG

60

## Patent Application US/07/864,692

```

521
522 GAGCCCGGAG CCCGGGTAGC GCGTAGAGCC GGC GCG ATG CAC GTG CGC TCA CTG      114
523                               Met His Val Arg Ser Leu
524                               1                    5
525
526 CGA GCT GCG GCG CCG CAC AGC TTC GTG GCG CTC TGG GCA CCC CTG TTC      162
527 Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala Pro Leu Phe
528           10                      15                      20
529
530 CTG CTG CGC TCC GCC CTG GCC GAC TTC AGC CTG GAC AAC GAG GTG CAC      210
531 Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn Glu Val His
532           25                      30                      35
533
534 TCG AGC TTC ATC CAC CGG CGC CTC CGC AGC CAG GAG CGG CGG GAG ATG      258
535 Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg Arg Glu Met
536           40                      45                      50
537
538 CAG CGC GAG ATC CTC TCC ATT TTG GGC TTG CCC CAC CGC CCG CGC CCG      306
539 Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg Pro Arg Pro
540           55                      60                      65                      70
541
542 CAC CTC CAG GGC AAG CAC AAC TCG GCA CCC ATG TTC ATG CTG GAC CTG      354
543 His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met Leu Asp Leu
544           75                      80                      85
545
546 TAC AAC GCC ATG GCG GTG GAG GAG GGC GGC GGG CCC GGC GGC CAG GGC      402
547 Tyr Asn Ala Met Ala Val Glu Glu Gly Gly Gly Pro Gly Gly Gln Gly
548           90                      95                      100
549
550 TTC TCC TAC CCC TAC AAG GCC GTC TTC AGT ACC CAG GGC CCC CCT CTG      450
551 Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly Pro Pro Leu
552           105                      110                      115
553
554 GCC AGC CTG CAA GAT AGC CAT TTC CTC ACC GAC GCC GAC ATG GTC ATG      498
555 Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp Met Val Met
556           120                      125                      130
557
558 AGC TTC GTC AAC CTC GTG GAA CAT GAC AAG GAA TTC TTC CAC CCA CGC      546
559 Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe His Pro Arg
560           135                      140                      145                      150
561
562 TAC CAC CAT CGA GAG TTC CGG TTT GAT CTT TCC AAG ATC CCA GAA GGG      594
563 Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile Pro Glu Gly
564           155                      160                      165
565
566 GAA GCT GTC ACG GCA GCC GAA TTC CGG ATC TAC AAG GAC TAC ATC CGG      642
567 Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp Tyr Ile Arg
568           170                      175                      180
569
570 GAA CGC TTC GAC AAT GAG ACG TTC CGG ATC AGC GTT TAT CAG GTG CTC      690
571 Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile Ser Val Tyr Gln Val Leu
572           185                      190                      195

```

## Patent Application US/07/864,692

```

573
574 CAG GAG CAC TTG GGC AGG GAA TCG GAT CTC TTC CTG CTC GAC AGC CGT      738
575 Gln Glu His Leu Gly Arg Glu Ser Asp Leu Phe Leu Leu Asp Ser Arg
576      200                                205                                210
577
578 ACC CTC TGG GCC TCG GAG GAG GGC TGG CTG GTG TTT GAC ATC ACA GCC      786
579 Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp Ile Thr Ala
580 215                                220                                225                                230
581
582 ACC AGC AAC CAC TGG GTG GTC AAT CCG CGG CAC AAC CTG GGC CTG CAG      834
583 Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu Gly Leu Gln
584                                235                                240                                245
585
586 CTC TCG GTG GAG ACG CTG GAT GGG CAG AGC ATC AAC CCC AAG TTG GCG      882
587 Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro Lys Leu Ala
588                                250                                255                                260
589
590 GGC CTG ATT GGG CGG CAC GGG CCC CAG AAC AAG CAG CCC TTC ATG GTG      930
591 Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro Phe Met Val
592                                265                                270                                275
593
594 GCT TTC TTC AAG GCC ACG GAG GTC CAC TTC CGC AGC ATC CGG TCC ACG      978
595 Ala Phe Phe Lys Ala Thr Glu Val His Phe Arg Ser Ile Arg Ser Thr
596                                280                                285                                290
597
598 GGG AGC AAA CAG CGC AGC CAG AAC CGC TCC AAG ACG CCC AAG AAC CAG      1026
599 Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro Lys Asn Gln
600 295                                300                                305                                310
601
602 GAA GCC CTG CGG ATG GCC AAC GTG GCA GAG AAC AGC AGC AGC GAC CAG      1074
603 Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser Ser Asp Gln
604                                315                                320                                325
605
606 AGG CAG GCC TGT AAG AAG CAC GAG CTG TAT GTC AGC TTC CGA GAC CTG      1122
607 Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu
608                                330                                335                                340
609
610 GGC TGG CAG GAC TGG ATC ATC GCG CCT GAA GGC TAC GCC GCC TAC TAC      1170
611 Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr Tyr
612                                345                                350                                355
613
614 TGT GAG GGG GAG TGT GCC TTC CCT CTG AAC TCC TAC ATG AAC GCC ACC      1218
615 Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala Thr
616 360                                365                                370
617
618 AAC CAC GCC ATC GTG CAG ACG CTG GTC CAC TTC ATC AAC CCG GAA ACG      1266
619 Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro Glu Thr
620 375                                380                                385                                390
621
622 GTG CCC AAG CCC TGC TGT GCG CCC ACG CAG CTC AAT GCC ATC TCC GTC      1314
623 Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser Val
624                                395                                400                                405

```

## Patent Application US/07/864,692

```

625
626 CTC TAC TTC GAT GAC AGC TCC AAC GTC ATC CTG AAG AAA TAC AGA AAC      1362
627 Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn
628           410                      415                      420
629
630 ATG GTG GTC CGG GCC TGT GGC TGC CAC TAGCTCCTCC GAGAATTCAG      1409
631 Met Val Val Arg Ala Cys Gly Cys His
632           425                      430
633
634 ACCCTTTGGG GCCAAGTTTT TCTGGATCCT CCATTGCTC      1448
635
636
637 (2) INFORMATION FOR SEQ ID NO:6:
638
639     (i) SEQUENCE CHARACTERISTICS:
640         (A) LENGTH: 431 amino acids
641         (B) TYPE: amino acid
642         (D) TOPOLOGY: linear
643
644     (ii) MOLECULE TYPE: protein
645
646     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
647
648 Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
649   1           5           10           15
650
651 Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
652           20           25           30
653
654 Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
655           35           40           45
656
657 Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
658           50           55           60
659
660 Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
661   65           70           75           80
662
663 Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly
664           85           90           95
665
666 Gly Pro Gly Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser
667           100          105          110
668
669 Thr Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr
670           115          120          125
671
672 Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys
673           130          135          140
674
675 Glu Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu
676   145          150          155          160

```

## Patent Application US/07/864,692

677  
678 Ser Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile  
679 165 170 175  
680  
681 Tyr Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile  
682 180 185 190  
683  
684 Ser Val Tyr Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu  
685 195 200 205  
686  
687 Phe Leu Leu Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu  
688 210 215 220  
689  
690 Val Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg  
691 225 230 235 240  
692  
693 His Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser  
694 245 250 255  
695  
696 Ile Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn  
697 260 265 270  
698  
699 Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe  
700 275 280 285  
701  
702 Arg Ser Ile Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser  
703 290 295 300  
704  
705 Lys Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu  
706 305 310 315 320  
707  
708 Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr  
709 325 330 335  
710  
711 Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu  
712 340 345 350  
713  
714 Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn  
715 355 360 365  
716  
717 Ser Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His  
718 370 375 380  
719  
720 Phe Ile Asn Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln  
721 385 390 395 400  
722  
723 Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile  
724 405 410 415  
725  
726 Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His  
727 420 425 430  
728

## Patent Application US/07/864,692

## 729 (2) INFORMATION FOR SEQ ID NO:7:

730

## 731 (i) SEQUENCE CHARACTERISTICS:

732 (A) LENGTH: 2923 base pairs

733 (B) TYPE: nucleic acid

734 (C) STRANDEDNESS: double

735 (D) TOPOLOGY: circular

736

737 (ii) MOLECULE TYPE: cDNA to mRNA

738

739 (iii) HYPOTHETICAL: NO

740

741 (vi) ORIGINAL SOURCE:

742 (A) ORGANISM: Homo sapiens

743 (F) TISSUE TYPE: Human placenta

744

745 (vii) IMMEDIATE SOURCE:

746 (A) LIBRARY: Stratagene catalog #936203 Human placenta  
747 cDNA library

748 (B) CLONE: BMP6C35

749

750 (viii) POSITION IN GENOME:

751 (C) UNITS: bp

752

753 (ix) FEATURE:

754 (A) NAME/KEY: CDS

755 (B) LOCATION: 160..1701

756

757 (ix) FEATURE:

758 (A) NAME/KEY: mat\_peptide

759 (B) LOCATION: 1282..1698

760

761 (ix) FEATURE:

762 (A) NAME/KEY: mRNA

763 (B) LOCATION: 1..2923

764

765

766 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

767

768 CGACCATGAG AGATAAGGAC TGAGGGCCAG GAAGGGGAAG CGAGCCCGCC GAGAGGTGGC 60

769

770 GGGGACTGCT CACGCCAAGG GCCACAGCGG CCGCGCTCCG GCCTCGCTCC GCCGCTCCAC 120

771

772 GCCTCGCGGG ATCCGCGGGG GCAGCCCGGC CGGGCGGGG ATG CCG GGG CTG GGG 174

773

Met Pro Gly Leu Gly

774

-374 -370

775

776 CGG AGG GCG CAG TGG CTG TGC TGG TGG TGG GGG CTG CTG TGC AGC TGC 222

777

Arg Arg Ala Gln Trp Leu Cys Trp Trp Trp Gly Leu Leu Cys Ser Cys

778

-365 -360 -355

779

780 TGC GGG CCC CCG CCG CTG CCG CCG CCC TTG CCC GCT GCC GCG GCC GCC 270

## Patent Application US/07/864,692

781	Cys Gly Pro Pro Pro Leu Arg Pro Pro Leu Pro Ala Ala Ala Ala Ala	
782	-350 -345 -340	
783		
784	GCC GCC GGG GGG CAG CTG CTG GGG GAC GGC GGG AGC CCC GGC CGC ACG	318
785	Ala Ala Gly Gly Gln Leu Leu Gly Asp Gly Gly Ser Pro Gly Arg Thr	
786	-335 -330 -325	
787		
788	GAG CAG CCG CCG CCG TCG CCG CAG TCC TCC TCG GGC TTC CTG TAC CGG	366
789	Glu Gln Pro Pro Pro Ser Pro Gln Ser Ser Ser Gly Phe Leu Tyr Arg	
790	-320 -315 -310	
791		
792	CGG CTC AAG ACG CAG GAG AAG CGG GAG ATG CAG AAG GAG ATC TTG TCG	414
793	Arg Leu Lys Thr Gln Glu Lys Arg Glu Met Gln Lys Glu Ile Leu Ser	
794	-305 -300 -295 -290	
795		
796	GTG CTG GGG CTC CCG CAC CGG CCC CGG CCC CTG CAC GGC CTC CAA CAG	462
797	Val Leu Gly Leu Pro His Arg Pro Arg Pro Leu His Gly Leu Gln Gln	
798	-285 -280 -275	
799		
800	CCG CAG CCC CCG GCG CTC CGG CAG CAG GAG GAG CAG CAG CAG CAG CAG	510
801	Pro Gln Pro Pro Ala Leu Arg Gln Gln Glu Glu Gln Gln Gln Gln Gln	
802	-270 -265 -260	
803		
804	CAG CTG CCT CGC GGA GAG CCC CCT CCC GGG CGA CTG AAG TCC GCG CCC	558
805	Gln Leu Pro Arg Gly Glu Pro Pro Pro Gly Arg Leu Lys Ser Ala Pro	
806	-255 -250 -245	
807		
808	CTC TTC ATG CTG GAT CTG TAC AAC GCC CTG TCC GCC GAC AAC GAC GAG	606
809	Leu Phe Met Leu Asp Leu Tyr Asn Ala Leu Ser Ala Asp Asn Asp Glu	
810	-240 -235 -230	
811		
812	GAC GGG GCG TCG GAG GGG GAG AGG CAG CAG TCC TGG CCC CAC GAA GCA	654
813	Asp Gly Ala Ser Glu Gly Glu Arg Gln Gln Ser Trp Pro His Glu Ala	
814	-225 -220 -215 -210	
815		
816	GCC AGC TCG TCC CAG CGT CGG CAG CCG CCC CCG GGC GCC GCG CAC CCG	702
817	Ala Ser Ser Ser Gln Arg Arg Gln Pro Pro Pro Gly Ala Ala His Pro	
818	-205 -200 -195	
819		
820	CTC AAC CGC AAG AGC CTT CTG GCC CCC GGA TCT GGC AGC GGC GGC GCG	750
821	Leu Asn Arg Lys Ser Leu Leu Ala Pro Gly Ser Gly Ser Gly Gly Ala	
822	-190 -185 -180	
823		
824	TCC CCA CTG ACC AGC GCG CAG GAC AGC GCC TTC CTC AAC GAC GCG GAC	798
825	Ser Pro Leu Thr Ser Ala Gln Asp Ser Ala Phe Leu Asn Asp Ala Asp	
826	-175 -170 -165	
827		
828	ATG GTC ATG AGC TTT GTG AAC CTG GTG GAG TAC GAC AAG GAG TTC TCC	846
829	Met Val Met Ser Phe Val Asn Leu Val Glu Tyr Asp Lys Glu Phe Ser	
830	-160 -155 -150	
831		
832	CCT CGT CAG CGA CAC CAC AAA GAG TTC AAG TTC AAC TTA TCC CAG ATT	894



## Patent Application US/07/864,692

833	Pro Arg Gln Arg His His Lys Glu Phe Lys Phe Asn Leu Ser Gln Ile	
834	-145 -140 -135 -130	
835		
836	CCT GAG GGT GAG GTG GTG ACG GCT GCA GAA TTC CGC ATC TAC AAG GAC	942
837	Pro Glu Gly Glu Val Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp	
838	-125 -120 -115	
839		
840	TGT GTT ATG GGG AGT TTT AAA AAC CAA ACT TTT CTT ATC AGC ATT TAT	990
841	Cys Val Met Gly Ser Phe Lys Asn Gln Thr Phe Leu Ile Ser Ile Tyr	
842	-110 -105 -100	
843		
844	CAA GTC TTA CAG GAG CAT CAG CAC AGA GAC TCT GAC CTG TTT TTG TTG	1038
845	Gln Val Leu Gln Glu His Gln His Arg Asp Ser Asp Leu Phe Leu Leu	
846	-95 -90 -85	
847		
848	GAC ACC CGT GTA GTA TGG GCC TCA GAA GAA GGC TGG CTG GAA TTT GAC	1086
849	Asp Thr Arg Val Val Trp Ala Ser Glu Glu Gly Trp Leu Glu Phe Asp	
850	-80 -75 -70	
851		
852	ATC ACG GCC ACT AGC AAT CTG TGG GTT GTG ACT CCA CAG CAT AAC ATG	1134
853	Ile Thr Ala Thr Ser Asn Leu Trp Val Val Thr Pro Gln His Asn Met	
854	-65 -60 -55 -50	
855		
856	GGG CTT CAG CTG AGC GTG GTG ACA AGG GAT GGA GTC CAC GTC CAC CCC	1182
857	Gly Leu Gln Leu Ser Val Val Thr Arg Asp Gly Val His Val His Pro	
858	-45 -40 -35	
859		
860	CGA GCC GCA GGC CTG GTG GGC AGA GAC GGC CCT TAC GAT AAG CAG CCC	1230
861	Arg Ala Ala Gly Leu Val Gly Arg Asp Gly Pro Tyr Asp Lys Gln Pro	
862	-30 -25 -20	
863		
864	TTC ATG GTG GCT TTC TTC AAA GTG AGT GAG GTC CAC GTG CGC ACC ACC	1278
865	Phe Met Val Ala Phe Phe Lys Val Ser Glu Val His Val Arg Thr Thr	
866	-15 -10 -5	
867		
868	AGG TCA GCC TCC AGC CGG CGC CGA CAA CAG AGT CGT AAT CGC TCT ACC	1326
869	Arg Ser Ala Ser Ser Arg Arg Arg Gln Gln Ser Arg Asn Arg Ser Thr	
870	1 5 10 15	
871		
872	CAG TCC CAG GAC GTG GCG CGG GTC TCC AGT GCT TCA GAT TAC AAC AGC	1374
873	Gln Ser Gln Asp Val Ala Arg Val Ser Ser Ala Ser Asp Tyr Asn Ser	
874	20 25 30	
875		
876	AGT GAA TTG AAA ACA GCC TGC AGG AAG CAT GAG CTG TAT GTG AGT TTC	1422
877	Ser Glu Leu Lys Thr Ala Cys Arg Lys His Glu Leu Tyr Val Ser Phe	
878	35 40 45	
879		
880	CAA GAC CTG GGA TGG CAG GAC TGG ATC ATT GCA CCC AAG GGC TAT GCT	1470
881	Gln Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala	
882	50 55 60	
883		
884	GCC AAT TAC TGT GAT GGA GAA TGC TCC TTC CCA CTC AAC GCA CAC ATG	1518

## Patent Application US/07/864,692

```
885  Ala Asn Tyr Cys Asp Gly Glu Cys Ser Phe Pro Leu Asn Ala His Met
886      65                      70                      75
887
888  AAT GCA ACC AAC CAC GCG ATT GTG CAG ACC TTG GTT CAC CTT ATG AAC      1566
889  Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Leu Met Asn
890      80                      85                      90                      95
891
892  CCC GAG TAT GTC CCC AAA CCG TGC TGT GCG CCA ACT AAG CTA AAT GCC      1614
893  Pro Glu Tyr Val Pro Lys Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala
894      100                      105                      110
895
896  ATC TCG GTT CTT TAC TTT GAT GAC AAC TCC AAT GTC ATT CTG AAA AAA      1662
897  Ile Ser Val Leu Tyr Phe Asp Asp Asn Ser Asn Val Ile Leu Lys Lys
898      115                      120                      125
899
900  TAC AGG AAT ATG GTT GTA AGA GCT TGT GGA TGC CAC TAACTCGAAA      1708
901  Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
902      130                      135                      140
903
904  CCAGATGCTG GGGACACACA TTCTGCCTTG GATTCCTAGA TTACATCTGC CTTAAAAAAA      1768
905
906  CACGGAAGCA CAGTTGGAGG TGGGACGATG AGACTTTGAA ACTATCTCAT GCCAGTGCCT      1828
907
908  TATTACCCAG GAAGATTTTA AAGGACCTCA TTAATAATTT GCTCACTTGG TAAATGACGT      1888
909
910  GAGTAGTTGT TGGTCTGTAG CAAGCTGAGT TTGGATGTCT GTAGCATAAG GTCTGGTAAC      1948
911
912  TGCAGAAACA TAACCGTGAA GCTCTTCCTA CCCTCCTCCC CCAAAAACCC ACCAAAATTA      2008
913
914  GTTTTAGCTG TAGATCAAGC TATTTGGGGT GTTTGTTAGT AAATAGGGAA AATAATCTCA      2068
915
916  AAGGAGTTAA ATGTATTCTT GGCTAAAGGA TCAGCTGGTT CAGTACTGTC TATCAAAGGT      2128
917
918  AGATTTTACA GAGAACAGAA ATCGGGGAAG TGGGGGGAAC GCCTCTGTTC AGTTCATTCC      2188
919
920  CAGAAGTCCA CAGGACGCAC AGCCCAGGCC ACAGCCAGGG CTCCACGGGG CGCCCTTGTC      2248
921
922  TCAGTCATTG CTGTTGTATG TTCGTGCTGG AGTTTTGTTG GTGTGAAAAT ACACTTATTT      2308
923
924  CAGCCAAAAC ATACCATTTC TACACCTCAA TCCTCCATTT GCTGTACTCT TTGCTAGTAC      2368
925
926  CAAAAGTAGA CTGATTACAC TGAGGTGAGG CTACAAGGGG TGTGTAACCG TGTAACACGT      2428
927
928  GAAGGCAGTG CTCACCTCTT CTTTACCAGA ACGGTTCTTT GACCAGCACA TTAACCTCTG      2488
929
930  GACTGCCGGC TCTAGTACCT TTTCAGTAAA GTGGTTCTCT GCCTTTTTTAC TATACAGCAT      2548
931
932  ACCACGCCAC AGGGTTAGAA CCAACGAAGA AAATAAAATG AGGGTGCCCA GCTTATAAGA      2608
933
934  ATGGTGTTAG GGGGATGAGC ATGCTGTTTA TGAACGAAA TCATGATTTC CCTGTAGAAA      2668
935
936  GTGAGGCTCA GATTAAATTT TAGAATATTT TCTAAATGTC TTTTTCACAA TCATGTGACT      2728
```

## Patent Application US/07/864,692

937-  
938 GGG AAGGCAA TTTCATACTA AACTGATTAA ATAATACATT TATAATCTAC AACTGTTTGC 2788  
939  
940 ACTTACAGCT TTTTGTGTA ATATAAACTA TAATTTATTG TCTATTTTAT ATCTGTTTTG 2848  
941  
942 CTGTGGCGTT GGGGGGGGGG CCGGGCTTTT GGGGGGGGGG GTTTGTTTGG GGGGTGTCGT 2908  
943  
944 GGTGTGGGCG GCGCG 2923  
945  
946  
947 (2) INFORMATION FOR SEQ ID NO:8:  
948  
949 (i) SEQUENCE CHARACTERISTICS:  
950 (A) LENGTH: 513 amino acids  
951 (B) TYPE: amino acid  
952 (D) TOPOLOGY: linear  
953  
954 (ii) MOLECULE TYPE: protein  
955  
956 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:  
957  
958 Met Pro Gly Leu Gly Arg Arg Ala Gln Trp Leu Cys Trp Trp Trp Gly  
959 -374 -370 -365 -360  
960  
961 Leu Leu Cys Ser Cys Cys Gly Pro Pro Pro Leu Arg Pro Pro Leu Pro  
962 -355 -350 -345  
963  
964 Ala Ala Ala Ala Ala Ala Ala Gly Gly Gln Leu Leu Gly Asp Gly Gly  
965 -340 -335 -330  
966  
967 Ser Pro Gly Arg Thr Glu Gln Pro Pro Pro Ser Pro Gln Ser Ser Ser  
968 -325 -320 -315  
969  
970 Gly Phe Leu Tyr Arg Arg Leu Lys Thr Gln Glu Lys Arg Glu Met Gln  
971 -310 -305 -300 -295  
972  
973 Lys Glu Ile Leu Ser Val Leu Gly Leu Pro His Arg Pro Arg Pro Leu  
974 -290 -285 -280  
975  
976 His Gly Leu Gln Gln Pro Gln Pro Pro Ala Leu Arg Gln Gln Glu Glu  
977 -275 -270 -265  
978  
979 Gln Gln Gln Gln Gln Gln Leu Pro Arg Gly Glu Pro Pro Pro Gly Arg  
980 -260 -255 -250  
981  
982 Leu Lys Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr Asn Ala Leu Ser  
983 -245 -240 -235  
984  
985 Ala Asp Asn Asp Glu Asp Gly Ala Ser Glu Gly Glu Arg Gln Gln Ser  
986 -230 -225 -220 -215  
987  
988 Trp Pro His Glu Ala Ala Ser Ser Ser Gln Arg Arg Gln Pro Pro Pro

## Patent Application US/07/864,692

```

989      -210      -205      -200
990
991 Gly Ala Ala His Pro Leu Asn Arg Lys Ser Leu Leu Ala Pro Gly Ser
992      -195      -190      -185
993
994 Gly Ser Gly Gly Ala Ser Pro Leu Thr Ser Ala Gln Asp Ser Ala Phe
995      -180      -175      -170
996
997 Leu Asn Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu Tyr
998      -165      -160      -155
999
1000 Asp Lys Glu Phe Ser Pro Arg Gln Arg His His Lys Glu Phe Lys Phe
1001 -150      -145      -140      -135
1002
1003 Asn Leu Ser Gln Ile Pro Glu Gly Glu Val Val Thr Ala Ala Glu Phe
1004      -130      -125      -120
1005
1006 Arg Ile Tyr Lys Asp Cys Val Met Gly Ser Phe Lys Asn Gln Thr Phe
1007      -115      -110      -105
1008
1009 Leu Ile Ser Ile Tyr Gln Val Leu Gln Glu His Gln His Arg Asp Ser
1010      -100      -95      -90
1011
1012 Asp Leu Phe Leu Leu Asp Thr Arg Val Val Trp Ala Ser Glu Glu Gly
1013      -85      -80      -75
1014
1015 Trp Leu Glu Phe Asp Ile Thr Ala Thr Ser Asn Leu Trp Val Val Thr
1016      -70      -65      -60      -55
1017
1018 Pro Gln His Asn Met Gly Leu Gln Leu Ser Val Val Thr Arg Asp Gly
1019      -50      -45      -40
1020
1021 Val His Val His Pro Arg Ala Ala Gly Leu Val Gly Arg Asp Gly Pro
1022      -35      -30      -25
1023
1024 Tyr Asp Lys Gln Pro Phe Met Val Ala Phe Phe Lys Val Ser Glu Val
1025      -20      -15      -10
1026
1027 His Val Arg Thr Thr Arg Ser Ala Ser Ser Arg Arg Arg Gln Gln Ser
1028      -5      1      5      10
1029
1030 Arg Asn Arg Ser Thr Gln Ser Gln Asp Val Ala Arg Val Ser Ser Ala
1031      15      20      25
1032
1033 Ser Asp Tyr Asn Ser Ser Glu Leu Lys Thr Ala Cys Arg Lys His Glu
1034      30      35      40
1035
1036 Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala
1037      45      50      55
1038
1039 Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly Glu Cys Ser Phe Pro
1040      60      65      70

```

## Patent Application US/07/864,692

1041  
1042 Leu Asn Ala His Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu  
1043 75 80 85 90  
1044  
1045 Val His Leu Met Asn Pro Glu Tyr Val Pro Lys Pro Cys Cys Ala Pro  
1046 95 100 105  
1047  
1048 Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Asn Ser Asn  
1049 110 115 120  
1050  
1051 Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys  
1052 125 130 135  
1053  
1054 His  
1055  
1056  
1057  
1058 (2) INFORMATION FOR SEQ ID NO:9:  
1059  
1060 (i) SEQUENCE CHARACTERISTICS:  
1061 (A) LENGTH: 2153 base pairs  
1062 (B) TYPE: nucleic acid  
1063 (C) STRANDEDNESS: double  
1064 (D) TOPOLOGY: linear  
1065  
1066 (iii) HYPOTHETICAL: NO  
1067  
1068 (vi) ORIGINAL SOURCE:  
1069 (A) ORGANISM: Homo sapiens  
1070 (H) CELL LINE: U2-OS osteosarcoma  
1071  
1072 (vii) IMMEDIATE SOURCE:  
1073 (A) LIBRARY: U2-OS human osteosarcoma cDNA library  
1074 (B) CLONE: U2-16  
1075  
1076 (viii) POSITION IN GENOME:  
1077 (C) UNITS: bp  
1078  
1079 (ix) FEATURE:  
1080 (A) NAME/KEY: CDS  
1081 (B) LOCATION: 699..2063  
1082  
1083 (ix) FEATURE:  
1084 (A) NAME/KEY: mat\_peptide  
1085 (B) LOCATION: 1647..2060  
1086  
1087 (ix) FEATURE:  
1088 (A) NAME/KEY: mRNA  
1089 (B) LOCATION: 1..2153  
1090  
1091  
1092 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

## Patent Application US/07/864,692

```

1093-
1094 CTGGTATATT TGTGCCTGCT GGAGGTGGAA TTAACAGTAA GAAGGAGAAA GGGATTGAAT      60
1095
1096 GGACTTACAG GAAGGATTTC AAGTAAATTC AGGGAAACAC ATTTACTTGA ATAGTACAAC      120
1097
1098 CTAGAGTATT ATTTTACACT AAGACGACAC AAAAGATGTT AAAGTTATCA CCAAGCTGCC      180
1099
1100 GGACAGATAT ATATTCCAAC ACCAAGGTGC AGATCAGCAT AGATCTGTGA TTCAGAAATC      240
1101
1102 AGGATTTGTT TTGGAAAGAG CTCAAGGGTT GAGAAGAACT CAAAAGCAAG TGAAGATTAC      300
1103
1104 TTTGGGAACT ACAGTTTATC AGAAGATCAA CTTTGTCTAA TTCAAATACC AAAGGCCTGA      360
1105
1106 TTATCATAAA TTCATATAGG AATGCATAGG TCATCTGATC AAATAATATT AGCCGTCTTC      420
1107
1108 TGCTACATCA ATGCAGCAAA AACTCTTAAC AACTGTGGAT AATTGGAAAT CTGAGTTTCA      480
1109
1110 GCTTTCTTAG AAATAACTAC TCTTGACATA TTCCAAAATA TTTAAAATAG GACAGGAAAA      540
1111
1112 TCGGTGAGGA TGTTGTGCTC AGAAATGTCA CTGTCATGAA AAATAGGTAA ATTTGTTTTT      600
1113
1114 TCAGCTACTG GGAAACTGTA CCTCCTAGAA CCTTAGGTTT TTTTTTTTTT AAGAGGACAA      660
1115
1116 GAAGGACTAA AAATATCAAC TTTTGCTTTT GGACAAAA ATG CAT CTG ACT GTA      713
1117                               Met His Leu Thr Val
1118                               -316-315
1119
1120 TTT TTA CTT AAG GGT ATT GTG GGT TTC CTC TGG AGC TGC TGG GTT CTA      761
1121 Phe Leu Leu Lys Gly Ile Val Gly Phe Leu Trp Ser Cys Trp Val Leu
1122   -310                      -305                      -300
1123
1124 GTG GGT TAT GCA AAA GGA GGT TTG GGA GAC AAT CAT GTT CAC TCC AGT      809
1125 Val Gly Tyr Ala Lys Gly Gly Leu Gly Asp Asn His Val His Ser Ser
1126 -295                      -290                      -285                      -280
1127
1128 TTT ATT TAT AGA AGA CTA CGG AAC CAC GAA AGA CGG GAA ATA CAA AGG      857
1129 Phe Ile Tyr Arg Arg Leu Arg Asn His Glu Arg Arg Glu Ile Gln Arg
1130                      -275                      -270                      -265
1131
1132 GAA ATT CTC TCT ATC TTG GGT TTG CCT CAC AGA CCC AGA CCA TTT TCA      905
1133 Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg Pro Arg Pro Phe Ser
1134                      -260                      -255                      -250
1135
1136 CCT GGA AAA ATG ACC AAT CAA GCG TCC TCT GCA CCT CTC TTT ATG CTG      953
1137 Pro Gly Lys Met Thr Asn Gln Ala Ser Ser Ala Pro Leu Phe Met Leu
1138 -245                      -240                      -235
1139
1140 GAT CTC TAC AAT GCC GAA GAA AAT CCT GAA GAG TCG GAG TAC TCA GTA      1001
1141 Asp Leu Tyr Asn Ala Glu Glu Asn Pro Glu Glu Ser Glu Tyr Ser Val
1142 -230                      -225                      -220
1143
1144 AGG GCA TCC TTG GCA GAA GAG ACC AGA GGG GCA AGA AAG GGA TAC CCA      1049

```

## Patent Application US/07/864,692

1145	Arg	Ala	Ser	Leu	Ala	Glu	Glu	Thr	Arg	Gly	Ala	Arg	Lys	Gly	Tyr	Pro	
1146	-215					-210					-205					-200	
1147																	
1148	GCC	TCT	CCC	AAT	GGG	TAT	CCT	CGT	CGC	ATA	CAG	TTA	TCT	CGG	ACG	ACT	1097
1149	Ala	Ser	Pro	Asn	Gly	Tyr	Pro	Arg	Arg	Ile	Gln	Leu	Ser	Arg	Thr	Thr	
1150				-195					-190						-185		
1151																	
1152	CCT	CTG	ACC	ACC	CAG	AGT	CCT	CCT	CTA	GCC	AGC	CTC	CAT	GAT	ACC	AAC	1145
1153	Pro	Leu	Thr	Thr	Gln	Ser	Pro	Pro	Leu	Ala	Ser	Leu	His	Asp	Thr	Asn	
1154			-180					-175					-170				
1155																	
1156	TTT	CTG	AAT	GAT	GCT	GAC	ATG	GTC	ATG	AGC	TTT	GTC	AAC	TTA	GTT	GAA	1193
1157	Phe	Leu	Asn	Asp	Ala	Asp	Met	Val	Met	Ser	Phe	Val	Asn	Leu	Val	Glu	
1158			-165					-160					-155				
1159																	
1160	AGA	GAC	AAG	GAT	TTT	TCT	CAC	CAG	CGA	AGG	CAT	TAC	AAA	GAA	TTT	CGA	1241
1161	Arg	Asp	Lys	Asp	Phe	Ser	His	Gln	Arg	Arg	His	Tyr	Lys	Glu	Phe	Arg	
1162	-150					-145					-140						
1163																	
1164	TTT	GAT	CTT	ACC	CAA	ATT	CCT	CAT	GGA	GAG	GCA	GTG	ACA	GCA	GCT	GAA	1289
1165	Phe	Asp	Leu	Thr	Gln	Ile	Pro	His	Gly	Glu	Ala	Val	Thr	Ala	Ala	Glu	
1166	-135				-130				-125				-120				
1167																	
1168	TTC	CGG	ATA	TAC	AAG	GAC	CGG	AGC	AAC	AAC	CGA	TTT	GAA	AAT	GAA	ACA	1337
1169	Phe	Arg	Ile	Tyr	Lys	Asp	Arg	Ser	Asn	Asn	Arg	Phe	Glu	Asn	Glu	Thr	
1170			-115					-110					-105				
1171																	
1172	ATT	AAG	ATT	AGC	ATA	TAT	CAA	ATC	ATC	AAG	GAA	TAC	ACA	AAT	AGG	GAT	1385
1173	Ile	Lys	Ile	Ser	Ile	Tyr	Gln	Ile	Ile	Lys	Glu	Tyr	Thr	Asn	Arg	Asp	
1174			-100					-95					-90				
1175																	
1176	GCA	GAT	CTG	TTC	TTG	TTA	GAC	ACA	AGA	AAG	GCC	CAA	GCT	TTA	GAT	GTG	1433
1177	Ala	Asp	Leu	Phe	Leu	Leu	Asp	Thr	Arg	Lys	Ala	Gln	Ala	Leu	Asp	Val	
1178			-85				-80					-75					
1179																	
1180	GGT	TGG	CTT	GTC	TTT	GAT	ATC	ACT	GTG	ACC	AGC	AAT	CAT	TGG	GTG	ATT	1481
1181	Gly	Trp	Leu	Val	Phe	Asp	Ile	Thr	Val	Thr	Ser	Asn	His	Trp	Val	Ile	
1182	-70					-65					-60						
1183																	
1184	AAT	CCC	CAG	AAT	AAT	TTG	GGC	TTA	CAG	CTC	TGT	GCA	GAA	ACA	GGG	GAT	1529
1185	Asn	Pro	Gln	Asn	Asn	Leu	Gly	Leu	Gln	Leu	Cys	Ala	Glu	Thr	Gly	Asp	
1186	-55				-50					-45					-40		
1187																	
1188	GGA	CGC	AGT	ATC	AAC	GTA	AAA	TCT	GCT	GGT	CTT	GTG	GGA	AGA	CAG	GGA	1577
1189	Gly	Arg	Ser	Ile	Asn	Val	Lys	Ser	Ala	Gly	Leu	Val	Gly	Arg	Gln	Gly	
1190			-35					-30					-25				
1191																	
1192	CCT	CAG	TCA	AAA	CAA	CCA	TTC	ATG	GTG	GCC	TTC	TTC	AAG	GCG	AGT	GAG	1625
1193	Pro	Gln	Ser	Lys	Gln	Pro	Phe	Met	Val	Ala	Phe	Phe	Lys	Ala	Ser	Glu	
1194			-20					-15					-10				
1195																	
1196	GTA	CTT	CTT	CGA	TCC	GTG	AGA	GCA	GCC	AAC	AAA	CGA	AAA	AAT	CAA	AAC	1673

## Patent Application US/07/864,692

1197 Val Leu Leu Arg Ser Val Arg Ala Ala Asn Lys Arg Lys Asn Gln Asn  
1198 -5 1 5  
1199  
1200 CGC AAT AAA TCC AGC TCT CAT CAG GAC TCC TCC AGA ATG TCC AGT GTT 1721  
1201 Arg Asn Lys Ser Ser Ser His Gln Asp Ser Ser Arg Met Ser Ser Val  
1202 10 15 20 25  
1203  
1204 GGA GAT TAT AAC ACA AGT GAG CAA AAA CAA GCC TGT AAG AAG CAC GAA 1769  
1205 Gly Asp Tyr Asn Thr Ser Glu Gln Lys Gln Ala Cys Lys Lys His Glu  
1206 30 35 40  
1207  
1208 CTC TAT GTG AGC TTC CGG GAT CTG GGA TGG CAG GAC TGG ATT ATA GCA 1817  
1209 Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala  
1210 45 50 55  
1211  
1212 CCA GAA GGA TAC GCT GCA TTT TAT TGT GAT GGA GAA TGT TCT TTT CCA 1865  
1213 Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp Gly Glu Cys Ser Phe Pro  
1214 60 65 70  
1215  
1216 CTT AAC GCC CAT ATG AAT GCC ACC AAC CAC GCT ATA GTT CAG ACT CTG 1913  
1217 Leu Asn Ala His Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu  
1218 75 80 85  
1219  
1220 GTT CAT CTG ATG TTT CCT GAC CAC GTA CCA AAG CCT TGT TGT GCT CCA 1961  
1221 Val His Leu Met Phe Pro Asp His Val Pro Lys Pro Cys Cys Ala Pro  
1222 90 95 100 105  
1223  
1224 ACC AAA TTA AAT GCC ATC TCT GTT CTG TAC TTT GAT GAC AGC TCC AAT 2009  
1225 Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn  
1226 110 115 120  
1227  
1228 GTC ATT TTG AAA AAA TAT AGA AAT ATG GTA GTA CGC TCA TGT GGC TGC 2057  
1229 Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ser Cys Gly Cys  
1230 125 130 135  
1231  
1232 CAC TAATATTAAA TAATATTGAT AATAACAAAA AGATCTGTAT TAAGGTTTAT 2110  
1233 His  
1234  
1235  
1236 GGCTGCAATA AAAAGCATAC TTTCAGACAA ACAGAAAAAA AAA 2153  
1237  
1238  
1239 (2) INFORMATION FOR SEQ ID NO:10:  
1240  
1241 (i) SEQUENCE CHARACTERISTICS:  
1242 (A) LENGTH: 454 amino acids  
1243 (B) TYPE: amino acid  
1244 (D) TOPOLOGY: linear  
1245  
1246 (ii) MOLECULE TYPE: protein  
1247  
1248 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:



## Patent Application US/07/864,692

1249-  
1250 Met His Leu Thr Val Phe Leu Leu Lys Gly Ile Val Gly Phe Leu Trp  
1251 -316 -315 -310 -305  
1252  
1253 Ser Cys Trp Val Leu Val Gly Tyr Ala Lys Gly Gly Leu Gly Asp Asn  
1254 -300 -295 -290 -285  
1255  
1256 His Val His Ser Ser Phe Ile Tyr Arg Arg Leu Arg Asn His Glu Arg  
1257 -280 -275 -270  
1258  
1259 Arg Glu Ile Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg  
1260 -265 -260 -255  
1261  
1262 Pro Arg Pro Phe Ser Pro Gly Lys Met Thr Asn Gln Ala Ser Ser Ala  
1263 -250 -245 -240  
1264  
1265 Pro Leu Phe Met Leu Asp Leu Tyr Asn Ala Glu Glu Asn Pro Glu Glu  
1266 -235 -230 -225  
1267  
1268 Ser Glu Tyr Ser Val Arg Ala Ser Leu Ala Glu Glu Thr Arg Gly Ala  
1269 -220 -215 -210 -205  
1270  
1271 Arg Lys Gly Tyr Pro Ala Ser Pro Asn Gly Tyr Pro Arg Arg Ile Gln  
1272 -200 -195 -190  
1273  
1274 Leu Ser Arg Thr Thr Pro Leu Thr Thr Gln Ser Pro Pro Leu Ala Ser  
1275 -185 -180 -175  
1276  
1277 Leu His Asp Thr Asn Phe Leu Asn Asp Ala Asp Met Val Met Ser Phe  
1278 -170 -165 -160  
1279  
1280 Val Asn Leu Val Glu Arg Asp Lys Asp Phe Ser His Gln Arg Arg His  
1281 -155 -150 -145  
1282  
1283 Tyr Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro His Gly Glu Ala  
1284 -140 -135 -130 -125  
1285  
1286 Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp Arg Ser Asn Asn Arg  
1287 -120 -115 -110  
1288  
1289 Phe Glu Asn Glu Thr Ile Lys Ile Ser Ile Tyr Gln Ile Ile Lys Glu  
1290 -105 -100 -95  
1291  
1292 Tyr Thr Asn Arg Asp Ala Asp Leu Phe Leu Leu Asp Thr Arg Lys Ala  
1293 -90 -85 -80  
1294  
1295 Gln Ala Leu Asp Val Gly Trp Leu Val Phe Asp Ile Thr Val Thr Ser  
1296 -75 -70 -65  
1297  
1298 Asn His Trp Val Ile Asn Pro Gln Asn Asn Leu Gly Leu Gln Leu Cys  
1299 -60 -55 -50 -45  
1300

## Patent Application US/07/864,692

1301 Ala Glu Thr Gly Asp Gly Arg Ser Ile Asn Val Lys Ser Ala Gly Leu  
1302 -40 -35 -30  
1303  
1304 Val Gly Arg Gln Gly Pro Gln Ser Lys Gln Pro Phe Met Val Ala Phe  
1305 -25 -20 -15  
1306  
1307 Phe Lys Ala Ser Glu Val Leu Leu Arg Ser Val Arg Ala Ala Asn Lys  
1308 -10 -5 1  
1309  
1310 Arg Lys Asn Gln Asn Arg Asn Lys Ser Ser Ser His Gln Asp Ser Ser  
1311 5 10 15 20  
1312  
1313 Arg Met Ser Ser Val Gly Asp Tyr Asn Thr Ser Glu Gln Lys Gln Ala  
1314 25 30 35  
1315  
1316 Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln  
1317 40 45 50  
1318  
1319 Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp Gly  
1320 55 60 65  
1321  
1322 Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala  
1323 70 75 80  
1324  
1325 Ile Val Gln Thr Leu Val His Leu Met Phe Pro Asp His Val Pro Lys  
1326 85 90 95 100  
1327  
1328 Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe  
1329 105 110 115  
1330  
1331 Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val  
1332 120 125 130  
1333  
1334 Arg Ser Cys Gly Cys His  
1335 135  
1336  
1337

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1003 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: cDNA to mRNA

## (iii) HYPOTHETICAL: NO

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(F) TISSUE TYPE: Human Heart

## Patent Application US/07/864,692

1353  
1354 (vii) IMMEDIATE SOURCE:  
1355 (A) LIBRARY: Human heart cDNA library stratagene catalog  
1356 #936208  
1357 (B) CLONE: hh38  
1358  
1359 (viii) POSITION IN GENOME:  
1360 (C) UNITS: bp  
1361  
1362 (ix) FEATURE:  
1363 (A) NAME/KEY: CDS  
1364 (B) LOCATION: 8..850  
1365  
1366 (ix) FEATURE:  
1367 (A) NAME/KEY: mat\_peptide  
1368 (B) LOCATION: 427..843  
1369  
1370 (ix) FEATURE:  
1371 (A) NAME/KEY: mRNA  
1372 (B) LOCATION: 1..997  
1373  
1374  
1375 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  
1376  
1377 GAATTCC GAG CCC CAT TGG AAG GAG TTC CGC TTT GAC CTG ACC CAG ATC 49  
1378 Glu Pro His Trp Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile  
1379 -139 -135 -130  
1380  
1381 CCG GCT GGG GAG GCG GTC ACA GCT GCG GAG TTC CGG ATT TAC AAG GTG 97  
1382 Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val  
1383 -125 -120 -115 -110  
1384  
1385 CCC AGC ATC CAC CTG CTC AAC AGG ACC CTC CAC GTC AGC ATG TTC CAG 145  
1386 Pro Ser Ile His Leu Leu Asn Arg Thr Leu His Val Ser Met Phe Gln  
1387 -105 -100 -95  
1388  
1389 GTG GTC CAG GAG CAG TCC AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT 193  
1390 Val Val Gln Glu Gln Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp  
1391 -90 -85 -80  
1392  
1393 CTT CAG ACG CTC CGA GCT GGA GAC GAG GGC TGG CTG GTG CTG GAT GTC 241  
1394 Leu Gln Thr Leu Arg Ala Gly Asp Glu Gly Trp Leu Val Leu Asp Val  
1395 -75 -70 -65  
1396  
1397 ACA GCA GCC AGT GAC TGC TGG TTG CTG AAG CGT CAC AAG GAC CTG GGA 289  
1398 Thr Ala Ala Ser Asp Cys Trp Leu Leu Lys Arg His Lys Asp Leu Gly  
1399 -60 -55 -50  
1400  
1401 CTC CGC CTC TAT GTG GAG ACT GAG GAT GGG CAC AGC GTG GAT CCT GGC 337  
1402 Leu Arg Leu Tyr Val Glu Thr Glu Asp Gly His Ser Val Asp Pro Gly  
1403 -45 -40 -35 -30  
1404

## Patent Application US/07/864,692

```
1405 CTG GCC GGC CTG CTG GGT CAA CGG GCC CCA CGC TCC CAA CAG CCT TTC 385
1406 Leu Ala Gly Leu Leu Gly Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe
1407 -25 -20 -15
1408
1409 GTG GTC ACT TTC TTC AGG GCC AGT CCG AGT CCC ATC CGC ACC CCT CGG 433
1410 Val Val Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg
1411 -10 -5 1
1412
1413 GCA GTG AGG CCA CTG AGG AGG AGG CAG CCG AAG AAA AGC AAC GAG CTG 481
1414 Ala Val Arg Pro Leu Arg Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu
1415 5 10 15
1416
1417 CCG CAG GCC AAC CGA CTC CCA GGG ATC TTT GAT GAC GTC CAC GGC TCC 529
1418 Pro Gln Ala Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His Gly Ser
1419 20 25 30 35
1420
1421 CAC GGC CGG CAG GTC TGC CGT CGG CAC GAG CTC TAC GTC AGC TTC CAG 577
1422 His Gly Arg Gln Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln
1423 40 45 50
1424
1425 GAC CTT GGC TGG CTG GAC TGG GTC ATC GCC CCC CAA GGC TAC TCA GCC 625
1426 Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala
1427 55 60 65
1428
1429 TAT TAC TGT GAG GGG GAG TGC TCC TTC CCG CTG GAC TCC TGC ATG AAC 673
1430 Tyr Tyr Cys Glu Gly Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn
1431 70 75 80
1432
1433 GCC ACC AAC CAC GCC ATC CTG CAG TCC CTG GTG CAC CTG ATG AAG CCA 721
1434 Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro
1435 85 90 95
1436
1437 AAC GCA GTC CCC AAG GCG TGC TGT GCA CCC ACC AAG CTG AGC GCC ACC 769
1438 Asn Ala Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr
1439 100 105 110 115
1440
1441 TCT GTG CTC TAC TAT GAC AGC AGC AAC AAC GTC ATC CTG CGC AAG CAC 817
1442 Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His
1443 120 125 130
1444
1445 CGC AAC ATG GTG GTC AAG GCC TGC GGC TGC CAC TGAGTCAGCC CGCCCAGCCC 870
1446 Arg Asn Met Val Val Lys Ala Cys Gly Cys His
1447 135 140
1448
1449 TACTGCAGCC ACCCTTCTCA TCTGGATCGG GCCCTGCAGA GGCAGAAAAC CCTTAAATGC 930
1450
1451 TGTCACAGCT CAAGCAGGAG TGTCAGGGGC CCTCACTCTC GGTGCCTACT TCCTGTCAGG 990
1452
1453 CTTCTGGGAA TTC 1003
1454
1455
1456 (2) INFORMATION FOR SEQ ID NO:12:
```

## Patent Application US/07/864,692

1457-  
1458 (i) SEQUENCE CHARACTERISTICS:  
1459 (A) LENGTH: 281 amino acids  
1460 (B) TYPE: amino acid  
1461 (D) TOPOLOGY: linear  
1462  
1463 (ii) MOLECULE TYPE: protein  
1464  
1465 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:  
1466  
1467 Glu Pro His Trp Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro Ala  
1468 -139 -135 -130 -125  
1469  
1470 Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val Pro Ser  
1471 -120 -115 -110  
1472  
1473 Ile His Leu Leu Asn Arg Thr Leu His Val Ser Met Phe Gln Val Val  
1474 -105 -100 -95  
1475  
1476 Gln Glu Gln Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln  
1477 -90 -85 -80  
1478  
1479 Thr Leu Arg Ala Gly Asp Glu Gly Trp Leu Val Leu Asp Val Thr Ala  
1480 -75 -70 -65 -60  
1481  
1482 Ala Ser Asp Cys Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu Arg  
1483 -55 -50 -45  
1484  
1485 Leu Tyr Val Glu Thr Glu Asp Gly His Ser Val Asp Pro Gly Leu Ala  
1486 -40 -35 -30  
1487  
1488 Gly Leu Leu Gly Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val  
1489 -25 -20 -15  
1490  
1491 Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg Ala Val  
1492 -10 -5 1 5  
1493  
1494 Arg Pro Leu Arg Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu Pro Gln  
1495 10 15 20  
1496  
1497 Ala Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His Gly Ser His Gly  
1498 25 30 35  
1499  
1500 Arg Gln Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu  
1501 40 45 50  
1502  
1503 Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr  
1504 55 60 65  
1505  
1506 Cys Glu Gly Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr  
1507 70 75 80 85  
1508

## Patent Application US/07/864,692

1509 Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala  
1510 90 95 100  
1511  
1512 Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val  
1513 105 110 115  
1514  
1515 Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn  
1516 120 125 130  
1517  
1518 Met Val Val Lys Ala Cys Gly Cys His  
1519 135 140  
1520

## (2) INFORMATION FOR SEQ ID NO:13:

1521  
1522  
1523 (i) SEQUENCE CHARACTERISTICS:  
1524 (A) LENGTH: 3623 base pairs  
1525 (B) TYPE: nucleic acid  
1526 (C) STRANDEDNESS: double  
1527 (D) TOPOLOGY: linear  
1528

1529 (ii) MOLECULE TYPE: DNA (genomic)  
1530  
1531

## (vii) IMMEDIATE SOURCE:

1532  
1533 (B) CLONE: pALBP2-781  
1534

## (ix) FEATURE:

1535 (A) NAME/KEY: CDS  
1536 (B) LOCATION: 2724..3071  
1537  
1538

## (ix) FEATURE:

1539 (A) NAME/KEY: terminator  
1540 (B) LOCATION: 3150..3218  
1541  
1542

## (ix) FEATURE:

1543 (A) NAME/KEY: RBS  
1544 (B) LOCATION: 2222..2723  
1545  
1546

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

1547  
1548  
1549  
1550 GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTTAA TGTCATGATA ATAATGGTTT 60  
1551  
1552 CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT 120  
1553  
1554 TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT 180  
1555  
1556 AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT ATCCCTTTT 240  
1557  
1558 TTGCGGCATT TTGCCTTCCT GTTTTGTCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG 300  
1559  
1560 CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA 360

## Patent Application US/07/864,692

1561  
1562 TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC 420  
1563  
1564 TATGTGGCGC GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC 480  
1565  
1566 ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG 540  
1567  
1568 GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA 600  
1569  
1570 ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTTG CACAACATGG 660  
1571  
1572 GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG 720  
1573  
1574 ACGAGCGTGA CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAA CTATTAAGT 780  
1575  
1576 GCGAACTACT TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG 840  
1577  
1578 TTGAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG 900  
1579  
1580 GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT 960  
1581  
1582 CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC 1020  
1583  
1584 AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT 1080  
1585  
1586 CATATATACT TTAGATTGAT TTAAAACTTC ATTTTAAATT TAAAAGGATC TAGGTGAAGA 1140  
1587  
1588 TCCTTTTTGA TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT 1200  
1589  
1590 CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT 1260  
1591  
1592 GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC 1320  
1593  
1594 TACCAACTCT TTTTCCGAAG GTAACGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC 1380  
1595  
1596 TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC 1440  
1597  
1598 TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG 1500  
1599  
1600 GGTGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGTT 1560  
1601  
1602 CGTGACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG 1620  
1603  
1604 AGCATTGAGA AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG 1680  
1605  
1606 GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT 1740  
1607  
1608 ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGA TGCTCGTCAG 1800  
1609  
1610 GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGGCCTT TTTACGGTTC CTGGCCTTTT 1860  
1611  
1612 GCTGGCCTTT TGCTCACATG TTCTTTCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA 1920

## Patent Application US/07/864,692

1613  
1614 TTACCGCCTT TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT 1980  
1615  
1616 CAGTGAGCGA GGAAGCGGAA GAGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC 2040  
1617  
1618 CGATTCATTA ATGCAGAATT GATCTCTCAC CTACCAAACA ATGCCCCCCT GCAAAAAATA 2100  
1619  
1620 AATTCATATA AAAAACATAC AGATAACCAT CTGCGGTGAT AAATTATCTC TGGCGGTGTT 2160  
1621  
1622 GACATAAATA CCACTGGCGG TGATACTGAG CACATCAGCA GGACGCACTG ACCACCATGA 2220  
1623  
1624 AGGTGACGCT CTTAAAAATT AAGCCCTGAA GAAGGGCAGC ATTCAAAGCA GAAGGCTTTG 2280  
1625  
1626 GGGTGTGTGA TACGAAACGA AGCATTGGCC GTAAGTGC GA TCCGGATTA GCTGCCAATG 2340  
1627  
1628 TGCCAATCGC GGGGGGTTTT CGTTCAGGAC TACAAC TGCC ACACACCACC AAAGCTAACT 2400  
1629  
1630 GACAGGAGAA TCCAGATGGA TGCACAAACA CGCCGCCGCG AACGTCGCGC AGAGAAACAG 2460  
1631  
1632 GCTCAATGGA AAGCAGCAAA TCCCCTGTTG GTTGGGGTAA GCGCAAAACC AGTTCCGAAA 2520  
1633  
1634 GATTTTTTTTA ACTATAAACG CTGATGGAAG CGTTTATGCG GAAGAGGTAA AGCCCTTCCC 2580  
1635  
1636 GAGTAACAAA AAAACAACAG CATAAATAAC CCCGCTCTTA CACATTCCAG CCCTGAAAAA 2640  
1637  
1638 GGGCATCAAA TTAAACCACA CCTATGGTGT ATGCATTTAT TTGCATACAT TCAATCAATT 2700  
1639  
1640 GTTATCTAAG GAAATACTTA CAT ATG CAA GCT AAA CAT AAA CAA CGT AAA 2750  
1641 Met Gln Ala Lys His Lys Gln Arg Lys  
1642 1 5  
1643  
1644 CGT CTG AAA TCT AGC TGT AAG AGA CAC CCT TTG TAC GTG GAC TTC AGT 2798  
1645 Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp Phe Ser  
1646 10 15 20 25  
1647  
1648 GAC GTG GGG TGG AAT GAC TGG ATT GTG GCT CCC CCG GGG TAT CAC GCC 2846  
1649 Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala  
1650 30 35 40  
1651  
1652 TTT TAC TGC CAC GGA GAA TGC CCT TTT CCT CTG GCT GAT CAT CTG AAC 2894  
1653 Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn  
1654 45 50 55  
1655  
1656 TCC ACT AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT GTT AAC TCT 2942  
1657 Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser  
1658 60 65 70  
1659  
1660 AAG ATT CCT AAG GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC TCG 2990  
1661 Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser  
1662 75 80 85  
1663  
1664 ATG CTG TAC CTT GAC GAG AAT GAA AAG GTT GTA TTA AAG AAC TAT CAG 3038



## Patent Application US/07/864,692

1665 Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn Tyr Gln  
1666 90 95 100 105  
1667  
1668 GAC ATG GTT GTG GAG GGT TGT GGG TGT CGC TAGTACAGCA AAATTAAATA 3088  
1669 Asp Met Val Val Glu Gly Cys Gly Cys Arg  
1670 110 115  
1671  
1672 CATAAATATA TATATATATA TATATTTTAG AAAAAAGAAA AAAATCTAGA GTCGACCTGC 3148  
1673  
1674 AGTAATCGTA CAGGGTAGTA CAAATAAAAA AGGCACGTCA GATGACGTGC CTTTTTCTT 3208  
1675  
1676 GTGAGCAGTA AGCTTGGCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCTGG 3268  
1677  
1678 CGTTACCCAA CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA 3328  
1679  
1680 AGAGGCCCCG ACCGATCGCC CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGCGCCT 3388  
1681  
1682 GATGCGGTAT TTTCTCCTTA CGCATCTGTG CGGTATTTCA CACCGCATAT ATGGTGCACT 3448  
1683  
1684 CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC 3508  
1685  
1686 GCTGACGCGC CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC 3568  
1687  
1688 GTCTCCGGGA GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGA 3623  
1689  
1690  
1691 (2) INFORMATION FOR SEQ ID NO:14:  
1692  
1693 (i) SEQUENCE CHARACTERISTICS:  
1694 (A) LENGTH: 115 amino acids  
1695 (B) TYPE: amino acid  
1696 (D) TOPOLOGY: linear  
1697  
1698 (ii) MOLECULE TYPE: protein  
1699  
1700 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:  
1701  
1702 Met Gln Ala Lys His Lys Gln Arg Lys Arg Leu Lys Ser Ser Cys Lys  
1703 1 5 10 15  
1704  
1705 Arg His Pro Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp  
1706 20 25 30  
1707  
1708 Ile Val Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly Glu Cys  
1709 35 40 45  
1710  
1711 Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Val  
1712 50 55 60  
1713  
1714 Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala Cys Cys  
1715 65 70 75 80  
1716

## Patent Application US/07/864,692

1717 Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu Asn  
1718 85 90 95  
1719

1720 Glu Lys Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu Gly Cys  
1721 100 105 110  
1722

1723 Gly Cys Arg  
1724 115  
1725  
1726

1727 (2) INFORMATION FOR SEQ ID NO:15:  
1728

1729 (i) SEQUENCE CHARACTERISTICS:  
1730 (A) LENGTH: 14 base pairs  
1731 (B) TYPE: nucleic acid  
1732 (C) STRANDEDNESS: single  
1733 (D) TOPOLOGY: linear  
1734

1735 (ii) MOLECULE TYPE: DNA (genomic)  
1736  
1737  
1738

1739 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:  
1740

1741 CATGGGCAGC TGAG  
1742

14

1743 (2) INFORMATION FOR SEQ ID NO:16:  
1744

1745 (i) SEQUENCE CHARACTERISTICS:  
1746 (A) LENGTH: 41 base pairs  
1747 (B) TYPE: nucleic acid  
1748 (C) STRANDEDNESS: single  
1749 (D) TOPOLOGY: linear  
1750

1751 (ii) MOLECULE TYPE: DNA (genomic)  
1752  
1753  
1754

1755 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:  
1756

1757 GAGGGTTGTG GGTGTCGCTA GTGAGTCGAC TACAGCAAAT T  
1758

41

1759 (2) INFORMATION FOR SEQ ID NO:17:  
1760

1761 (i) SEQUENCE CHARACTERISTICS:  
1762 (A) LENGTH: 38 base pairs  
1763 (B) TYPE: nucleic acid  
1764 (C) STRANDEDNESS: single  
1765 (D) TOPOLOGY: linear  
1766

1767 (ii) MOLECULE TYPE: DNA (genomic)  
1768

## Patent Application US/07/864,692

1769  
1770  
1771 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:  
1772  
1773 GGATGTGGGT GCCGCTGACT CTAGAGTCGA CGGAATTC 38  
1774  
1775 (2) INFORMATION FOR SEQ ID NO:18:  
1776  
1777 (i) SEQUENCE CHARACTERISTICS:  
1778 (A) LENGTH: 31 base pairs  
1779 (B) TYPE: nucleic acid  
1780 (C) STRANDEDNESS: single  
1781 (D) TOPOLOGY: linear  
1782  
1783 (ii) MOLECULE TYPE: DNA (genomic)  
1784  
1785  
1786  
1787 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:  
1788  
1789 AATTCACCAT GATTCCTGGT AACCGAATGC T 31  
1790  
1791 (2) INFORMATION FOR SEQ ID NO:19:  
1792  
1793 (i) SEQUENCE CHARACTERISTICS:  
1794 (A) LENGTH: 25 base pairs  
1795 (B) TYPE: nucleic acid  
1796 (C) STRANDEDNESS: single  
1797 (D) TOPOLOGY: linear  
1798  
1799 (ii) MOLECULE TYPE: DNA (genomic)  
1800  
1801  
1802  
1803 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:  
1804  
1805 GTGGTACTAA GGACCATTTGG CTTAC 25  
1806  
1807 (2) INFORMATION FOR SEQ ID NO:20:  
1808  
1809 (i) SEQUENCE CHARACTERISTICS:  
1810 (A) LENGTH: 27 base pairs  
1811 (B) TYPE: nucleic acid  
1812 (C) STRANDEDNESS: single  
1813 (D) TOPOLOGY: linear  
1814  
1815 (ii) MOLECULE TYPE: DNA (genomic)  
1816  
1817  
1818  
1819 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:  
1820

## Patent Application US/07/864,692

1821 CGACCTGCAG CCATGCATCT GACTGTA 27  
1822  
1823 (2) INFORMATION FOR SEQ ID NO:21:  
1824  
1825 (i) SEQUENCE CHARACTERISTICS:  
1826 (A) LENGTH: 27 base pairs  
1827 (B) TYPE: nucleic acid  
1828 (C) STRANDEDNESS: single  
1829 (D) TOPOLOGY: linear  
1830  
1831 (ii) MOLECULE TYPE: DNA (genomic)  
1832  
1833  
1834  
1835 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:  
1836  
1837 TGCCTGCAGT TTAATATTAG TGGCAGC 27  
1838  
1839 (2) INFORMATION FOR SEQ ID NO:22:  
1840  
1841 (i) SEQUENCE CHARACTERISTICS:  
1842 (A) LENGTH: 15 base pairs  
1843 (B) TYPE: nucleic acid  
1844 (C) STRANDEDNESS: single  
1845 (D) TOPOLOGY: linear  
1846  
1847 (ii) MOLECULE TYPE: DNA (genomic)  
1848  
1849  
1850  
1851 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:  
1852  
1853 CGACCTGCAG CCACC 15  
1854  
1855 (2) INFORMATION FOR SEQ ID NO:23:  
1856  
1857 (i) SEQUENCE CHARACTERISTICS:  
1858 (A) LENGTH: 81 base pairs  
1859 (B) TYPE: nucleic acid  
1860 (C) STRANDEDNESS: single  
1861 (D) TOPOLOGY: linear  
1862  
1863 (ii) MOLECULE TYPE: DNA (genomic)  
1864  
1865  
1866  
1867 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:  
1868  
1869 TCGACCCACC ATGCCGGGGC TGGGGCGGAG GGCGCAGTGG CTGTGCTGGT GGTGGGGGCT 60  
1870  
1871 GTGCTGCAGC TGCTGCGGGC C 81  
1872

## Patent Application US/07/864,692

1873 (2) INFORMATION FOR SEQ ID NO:24:  
1874

1875 (i) SEQUENCE CHARACTERISTICS:

1876 (A) LENGTH: 73 base pairs

1877 (B) TYPE: nucleic acid

1878 (C) STRANDEDNESS: single

1879 (D) TOPOLOGY: linear  
18801881 (ii) MOLECULE TYPE: DNA (genomic)  
1882

1883

1884

1885 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:  
1886

1887 CGCAGCAGCT GCACAGCAGC CCCACCACC AGCACAGCCA CTGCGCCCTC CGCCCCAGCC 60

1888

1889 CCGGCATGGT GGG 73  
18901891 (2) INFORMATION FOR SEQ ID NO:25:  
1892

1893 (i) SEQUENCE CHARACTERISTICS:

1894 (A) LENGTH: 11 base pairs

1895 (B) TYPE: nucleic acid

1896 (C) STRANDEDNESS: single

1897 (D) TOPOLOGY: linear  
18981899 (ii) MOLECULE TYPE: DNA (genomic)  
1900

1901

1902

1903 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:  
19041905 TCGACTGGTT T 11  
19061907 (2) INFORMATION FOR SEQ ID NO:26:  
1908

1909 (i) SEQUENCE CHARACTERISTICS:

1910 (A) LENGTH: 9 base pairs

1911 (B) TYPE: nucleic acid

1912 (C) STRANDEDNESS: single

1913 (D) TOPOLOGY: linear  
19141915 (ii) MOLECULE TYPE: DNA (genomic)  
1916

1917

1918

1919 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:  
19201921 CGAAACCAG 9  
19221923 (2) INFORMATION FOR SEQ ID NO:27:  
1924

## Patent Application US/07/864,692

1925 (i) SEQUENCE CHARACTERISTICS:  
1926 (A) LENGTH: 18 base pairs  
1927 (B) TYPE: nucleic acid  
1928 (C) STRANDEDNESS: single  
1929 (D) TOPOLOGY: linear  
1930  
1931 (ii) MOLECULE TYPE: DNA (genomic)  
1932  
1933  
1934  
1935 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:  
1936  
1937 TCGACAGGCT CGCCTGCA 18  
1938  
1939 (2) INFORMATION FOR SEQ ID NO:28:  
1940  
1941 (i) SEQUENCE CHARACTERISTICS:  
1942 (A) LENGTH: 10 base pairs  
1943 (B) TYPE: nucleic acid  
1944 (C) STRANDEDNESS: single  
1945 (D) TOPOLOGY: linear  
1946  
1947 (ii) MOLECULE TYPE: DNA (genomic)  
1948  
1949  
1950  
1951 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:  
1952  
1953 GTCCGAGCGG 10  
1954  
1955 (2) INFORMATION FOR SEQ ID NO:29:  
1956  
1957 (i) SEQUENCE CHARACTERISTICS:  
1958 (A) LENGTH: 29 base pairs  
1959 (B) TYPE: nucleic acid  
1960 (C) STRANDEDNESS: single  
1961 (D) TOPOLOGY: linear  
1962  
1963 (ii) MOLECULE TYPE: DNA (genomic)  
1964  
1965  
1966  
1967 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:  
1968  
1969 CAGGTCGACC CACCATGCAC GTGCGCTCA 29  
1970  
1971 (2) INFORMATION FOR SEQ ID NO:30:  
1972  
1973 (i) SEQUENCE CHARACTERISTICS:  
1974 (A) LENGTH: 27 base pairs  
1975 (B) TYPE: nucleic acid  
1976 (C) STRANDEDNESS: single

## Patent Application US/07/864,692

1977 (D) TOPOLOGY: linear  
1978  
1979 (ii) MOLECULE TYPE: DNA (genomic)  
1980  
1981  
1982  
1983 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:  
1984  
1985 TCTGTCGACC TCGGAGGAGC TAGTGGC

27

PAGE: 1

SEQUENCE VERIFICATION REPORT  
PATENT APPLICATION US/07/ ,692

DATE: 12/01/92  
TIME: 12:42:13  
S4120

LINE ERROR

ORIGINAL TEXT

29 Wrong application Serial Number  
747 Response Exceeds Line Limitations  
1356 Response Exceeds Line Limitations

(A) APPLICATION NUMBER: US  
cDNA library  
#936208





PAGE: 1

SEQUENCE MISSING ITEM REPORT  
PATENT APPLICATION US/07/4,692

DATE: 12/01/92  
TIME: 12:42:13  
S4120

MANDATORY IDENTIFIER THAT WAS NOT FOUND

PRIOR APPLICATION DATA  
APPLICATION NUMBER  
FILING DATE

PAGE: 1

SEQUENCE CORRECTION REPORT  
PATENT APPLICATION US/07/ ,692

DATE: 12/01/92  
TIME: 12:42:13  
S4120

LINE ORIGINAL TEXT

CORRECTED TEXT